

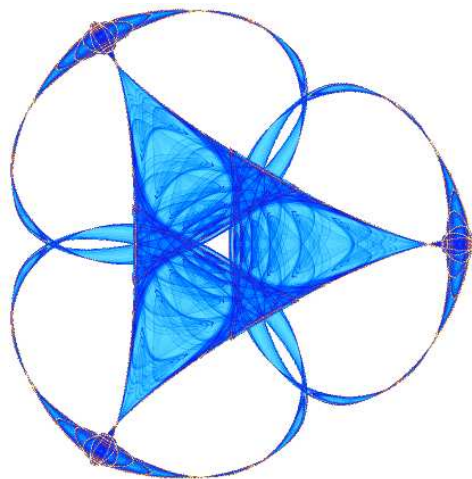
**IMAGE MODELING AND ENHANCEMENT VIA  
STRUCTURED SPARSE MODEL SELECTION**

By

**Guoshen Yu**  
**Guillermo Sapiro**  
and  
**Stéphane Mallat**

**IMA Preprint Series # 2292**

(January 2010)



**INSTITUTE FOR MATHEMATICS AND ITS APPLICATIONS**

UNIVERSITY OF MINNESOTA  
400 Lind Hall  
207 Church Street S.E.  
Minneapolis, Minnesota 55455-0436  
Phone: 612/624-6066 Fax: 612/626-7370  
URL: <http://www.ima.umn.edu>

Report Documentation Page				Form Approved OMB No. 0704-0188	
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE <b>JAN 2010</b>		2. REPORT TYPE		3. DATES COVERED <b>00-00-2010 to 00-00-2010</b>	
4. TITLE AND SUBTITLE <b>Image Modeling and Enhancement via Structured Sparse Model Selection</b>				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) <b>University of Minnesota, Institute for Mathematics and Its Applications, Minneapolis, MN, 55455-0436</b>				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT <b>Approved for public release; distribution unlimited</b>					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT <b>see report</b>					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT <b>Same as Report (SAR)</b>	18. NUMBER OF PAGES <b>5</b>	19a. NAME OF RESPONSIBLE PERSON
a. REPORT <b>unclassified</b>	b. ABSTRACT <b>unclassified</b>	c. THIS PAGE <b>unclassified</b>			

# IMAGE MODELING AND ENHANCEMENT VIA STRUCTURED SPARSE MODEL SELECTION

Guoshen Yu<sup>†</sup>, Guillermo Sapiro<sup>†</sup>, and Stéphane Mallat<sup>‡</sup>

<sup>†</sup>ECE, University of Minnesota    <sup>‡</sup>CMAP, Ecole Polytechnique

## ABSTRACT

An image representation framework based on structured sparse model selection is introduced in this work. The corresponding modeling dictionary is comprised of a family of learned orthogonal bases. For an image patch, a model is first selected from this dictionary through linear approximation in a best basis, and the signal estimation is then calculated with the selected model. The model selection leads to a guaranteed near optimal denoising estimator. The degree of freedom in the model selection is equal to the number of the bases, typically about 10 for natural images, and is significantly lower than with traditional overcomplete dictionary approaches, stabilizing the representation. For an image patch of size  $\sqrt{N} \times \sqrt{N}$ , the computational complexity of the proposed framework is  $\mathcal{O}(N^2)$ , typically 2 to 3 orders of magnitude faster than estimation in an overcomplete dictionary. The orthogonal bases are adapted to the image of interest and are computed with a simple and fast procedure. State-of-the-art results are shown in image denoising, deblurring, and inpainting.

**Index Terms**— Model selection, structured sparsity, best basis, denoising, deblurring, inpainting

## 1. INTRODUCTION

Image enhancement problems, such as denoising, deblurring and inpainting, are typical and important tasks in image processing. Estimation in sparse representations provides a powerful tool for such image enhancement tasks. State-of-the-art image enhancement results have been recently obtained with learned overcomplete dictionaries and sparse coding over them [1, 6, 14]. In these approaches, overcomplete local dictionaries are learned from the images, and the image enhancement is performed through estimation (sparse coding) in the resulting overcomplete sparse image representations.

Sparse signal estimation in an overcomplete dictionary is subject to an exponentially large degree of freedom. Calculating such estimation is not only computationally demanding, the resulting estimate is also likely to be unstable [2, 8], which may deteriorate the image enhancement results [15].

In this paper, we introduce a novel image representation framework, namely structured sparse model selection (SSMS). The dictionary is comprised of a family of learned orthogonal bases. For a given image patch, a model is selected through linear approximation in a best basis, over which the signal estimation is then calculated. The model selection leads to a guaranteed near optimal denoising estimator. The degree of freedom in the model selection is equal to the number of the bases, typically about 10 for natural images, significantly reduced with respect to the traditional overcomplete dictionary approaches, and thereby stabilizing the representation. For an image patch of size  $\sqrt{N} \times \sqrt{N}$ , signal estimation with SSMS is calculated with a complexity  $\mathcal{O}(KMN)$ , where  $K$  and  $M$  are respectively the number of the bases and the linear approximation dimension, whose typical values lead to a complexity in  $\mathcal{O}(N^2)$ , 2 to 3 orders of magnitude faster than estimation in an overcomplete dictionary. The orthogonal bases are adapted to the image of interest

and are computed with a simple and fast procedure. State-of-the-art results are shown in image denoising, deblurring and inpainting.

## 2. STATE-OF-THE-ART ON IMAGE ENHANCEMENT

State-of-the-art image enhancement is obtained by sparse coding image patches with learned overcomplete dictionaries [6, 14]. Images are decomposed into overlapping patches  $f \in \mathbb{R}^N$ , where  $\sqrt{N} \times \sqrt{N}$  is the patch size. A patch is assumed to be sparsely represented in an overcomplete dictionary  $\mathcal{D} = \{\phi_m\}_{m \in \Gamma}$ , with  $|\Gamma| \geq N$ :

$$f = f_\Lambda + w_\Lambda \quad \text{with} \quad f_\Lambda = \sum_{m \in \Lambda} a[m] \phi_m, \quad (1)$$

where the approximation error  $\|w_\Lambda\|^2 \ll \|f\|^2$  and  $|\Lambda| \ll |\Gamma|$ . A sparse approximation  $\tilde{f} = \sum_{m \in \tilde{\Lambda}} \tilde{a}[m] \phi_m$  of  $f$  is calculated with, for example, an orthogonal matching pursuit that approximates an  $l^0$  solution [18] or with an  $l^1$  minimization [5] that calculates

$$\tilde{a} = \arg \min_a \|f - \sum_{m \in \Gamma} a[m] \phi_m\|^2 + \lambda \|a\|_1. \quad (2)$$

Given an image patch data set  $\{f_i\}_{1 \leq i \leq I}$  with  $I \gg N$ , the dictionary  $\mathcal{D} = \{\phi_m\}_{m \in \Gamma}$  is learned by solving the following minimization, so that it provides sparse representations for all the image patches,

$$\min_{\{\phi_m\}_{m \in \Gamma}, \{a_i\}_{i=1, \dots, I}} \sum_{i=1}^I \|f_i - \sum_{m \in \Gamma} a_i[m] \phi_m\|^2 + \lambda \|a_i\|_1. \quad (3)$$

The minimization (3) is typically non-convex and a local solution is calculated with an iterative procedure, e.g., the K-SVD [1].



**Fig. 1.** Sparse signal models. Left: the overcomplete dictionary model. In this model any sparse set of atoms (columns in gray) can be selected. Right: the structured sparse model selection model. In this model, the dictionary is comprised of a family of orthogonal bases, only one selected at a time. The model selection is calculated with linear approximation in a best basis, i.e., only the first few atoms of one basis (in gray) can be selected.

## 3. STRUCTURED SPARSE MODEL SELECTION

The underlying signal model of the overcomplete dictionary approach, as depicted in Fig. 1, is that an image patch  $f \in \mathbb{R}^N$  can be accurately approximated with a few atoms  $\{\phi_m\}_{m \in \Lambda}$  in an overcomplete dictionary  $\mathcal{D} = \{\phi_m\}_{m \in \Gamma}$  with  $|\Gamma| \geq N \gg |\Lambda|$ . To approximate  $f$  with  $|\Lambda|$  atoms in the dictionary, the degree of freedom  $\binom{|\Gamma|}{|\Lambda|}$  in selecting the atoms is exponentially large ( $\sim 10^{14}$  for typical values  $N = 64$ ,  $|\Gamma| = 256$ ,  $|\Lambda| = 8$ ), which may lead to unstable signal estimation [2]. It has been shown that the  $l^1$  minimization achieves more stable recovery of the support  $\tilde{\Lambda}$  than  $l^0$  [13]. However, reliable signal estimation requires further stabilization [8, 15]. Attempting to exploit all the possibilities, the approximation (1) is computed via an iterative  $l^1$  minimization procedure (2), that has a relatively high computational complexity,  $\mathcal{O}(L|\Gamma|N)$ , where  $L$  is the number of iterations typically in the order of  $10^2$  to  $10^3$ .

Putting structure in sparsity stabilizes the estimation [8]. In this section, we introduce a structured sparse model selection (SSMS) signal model that reduces tremendously the degree of freedom as well as the computational complexity in signal estimation, and on the top of that, provides state-of-the-art image enhancement results.

### 3.1. Structured sparse model

As illustrated in Fig. 1, SSMS is defined with a dictionary comprised of  $K$  sub-dictionaries  $\mathcal{D} = \{\mathcal{B}^k\}_{1 \leq k \leq K}$ , each  $\mathcal{B}^k = \{\phi_m^k\}_{1 \leq m \leq N}$  being an orthogonal basis. An image patch  $f \in \mathbb{R}^N$  is assumed to be accurately approximated with a linear approximation in a best orthogonal basis indexed by  $k_0$

$$f = f_M^{k_0} + w_M^{k_0} \text{ with } f_M^{k_0} = \sum_{m=1}^M \langle f, \phi_m^{k_0} \rangle \phi_m^{k_0}, \quad (4)$$

where the linear approximation error  $\|w_M^{k_0}\|^2 \ll \|f\|^2$  and  $M \ll N$  is the linear approximation dimension.

Compared with the overcomplete dictionary model (1), the SSMS model (4) is much more structured and the structure comes from two concepts. First, the estimation in the best orthogonal basis excludes the possibility to select atoms across the bases, which shrinks the degree of freedom in atom selection from  $\binom{KN}{M}$  to  $K \binom{N}{M}$ . Second, the linear approximation in each basis further reduces the degree of freedom to just  $K$ , stabilizing the encoding and reducing the computational cost.

### 3.2. Structured model selection

The best basis  $\mathcal{B}^{k_0}$  in (4) is computed by maximizing the projection energy on  $\{\phi_m^k\}_{1 \leq m \leq M}$  over all the orthogonal bases  $\mathcal{B}^k \in \mathcal{D}$ :

$$k_0 = \arg \max_{k=1, \dots, K} \|f_M^k\|^2 = \arg \max_{k=1, \dots, K} \sum_{m=1}^M |\langle f, \phi_m^k \rangle|^2. \quad (5)$$

The model selection (5) has a degree of freedom  $K$  and is calculated with a complexity  $\mathcal{O}(KMN)$ . With the typical values  $N = 64$ ,  $M = 8$ ,  $K = 18$  as we will see in Section 3.3, the model selection complexity is  $\mathcal{O}(N^2)$ .

The following theorem shows that when the true original signal is contaminated by a Gaussian white noise  $w$ ,

$$y[n] = f[n] + w[n] \text{ for } 1 \leq n \leq N, \quad (6)$$

the empirical model selection leads to a near optimal denoising estimator whose risk is upper bounded by 4 times the best estimation obtained with an oracle.

**Theorem 1.** Let  $\sigma^2$  be the noise variance and  $T = \lambda \sigma \sqrt{\log_e(KN)}$  with  $\lambda \geq \sqrt{32 + \frac{8}{\log_e(KN)}}$ . For any  $f \in \mathbb{R}^N$ , the empirical model selection

$$k_0 = \arg \max_{k=1, \dots, K} \|y_M^k\|^2 = \arg \max_{k=1, \dots, K} \sum_{m=1}^M |\langle y, \phi_m^k \rangle|^2 \quad (7)$$

yields a linear projection estimator

$$\tilde{f} = \sum_{m=1}^M \langle y, \phi_m^{k_0} \rangle \phi_m^{k_0}, \quad (8)$$

of  $f$ , which satisfies

$$E[\|\tilde{f} - f\|^2] \leq 4 \min_{k=1, \dots, K} (\|f - f_M^k\|^2 + T^2 M) + \frac{32\sigma^2}{KN}, \quad (9)$$

where  $f_M^k = \sum_{m=1}^M \langle f, \phi_m^k \rangle \phi_m^k$ .

Theorem 1 is derived from the general model selection result of Barron, Birgé and Massart [3], which shows that the projector estimator over the best empirical set in an overcomplete dictionary is near optimal. Assuming linear approximation in a best orthogonal basis, the structured sparse model selection (7) is computed calculating the linear projection energy in each basis, and enjoys the advantage of being free of the thresholding parameter  $T$ , on contrast to [3].

### 3.3. Bases design

The orthogonal bases  $\{\mathcal{B}^k\}_{1 \leq k \leq K}$  are initialized with Principal Components Analysis (PCA) over synthetic contour images. Based on the model selection, they are then adapted to the image of interest by applying PCAs again over that image.

#### 3.3.1. Directional bases initialization



Fig. 2. The first 8 atoms in  $\mathcal{B}^k$  with  $\theta_k = 30^\circ$ .

As shown in [1], the most prominent atoms in a learned overcomplete dictionary look like edge patterns. In order to capture directional patterns, the orthogonal bases in SSMS are initialized by computing PCAs over synthetic contour images of various orientations, one basis for each orientation. At the orientation  $\theta_k$ , synthetic contours with different amount of blur are created. Image patches  $\{f_i\}_{i=1, \dots, I}$  of size  $\sqrt{N} \times \sqrt{N}$  touching the contours at different locations are collected. The orthogonal basis  $\mathcal{B}^k$  is obtained by calculating PCA over these patches.

The eigenvalues of the dictionary atoms decrease rapidly and become negligible after about the  $\sqrt{N}$ -th atom. As shown in Fig. 2, the first about  $\sqrt{N} = 8$  atoms oscillate in the direction orthogonal to  $\theta_k$  from low- to high-frequency and are quasi-constant in  $\theta_k$ . Other atoms that complete the orthogonal basis look unmeaningful. The orientation  $\theta_k$  is uniformly sampled from  $0^\circ$  to  $180^\circ$ , typically with a step of  $10^\circ$ , which results in  $K = 18$  orthogonal bases.

In order to better capture more complex texture patterns, a DCT basis is added to the dictionary  $\mathcal{D}$ .

#### 3.3.2. Bases adapted to image

The bases are adapted to the image of interest by applying the PCAs over the image patches, grouped following the model selection. The structured model selection with the directional bases introduced above assigns a model  $k$  to each patch. The basis  $\mathcal{B}^k$ , with  $1 \leq k \leq K$ , is adapted to the image by recalculating the PCA with all the image patches that have been assigned to the model  $k$ .

### 3.4. Best basis estimation

The signal estimation is calculated with a non-linear thresholding estimator in the best basis  $\mathcal{B}^{k_0}$

$$\tilde{f} = \sum_{m \in \Lambda} \langle y, \phi_m^{k_0} \rangle \phi_m^{k_0}, \quad (10)$$

where  $\Lambda = \{m : |\langle y, \phi_m^{k_0} \rangle| > T\}$ . In effect, the non-linear estimation typically improves the estimate with respect to the linear estimator (8), as the former is more robust to the signal model error.

## 4. EXPERIMENTAL RESULTS & CONCLUSIONS

### 4.1. Denoising

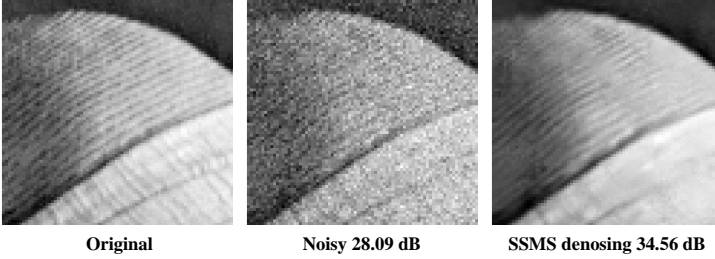
#### 4.1.1. Gaussian white noise removal

We first tested on images corrupted with Gaussian white noise of variance  $\sigma^2$ . The SSMS denoising is conducted with patches of size  $8 \times 8$ ,  $M = 8$ ,  $K = 18$  and  $T = 3\sigma$ . Table 1 summarizes its performance in comparison with the learned overcomplete dictionary (LOD) [6], the Non-Local Means (NLmeans) [4], and the Gaussian Scale Mixtures (GSM) [17]. The results of LOD and GSM are quoted from [6, 17], and those of NLmeans are produced with the code from Buades [4]. SSMS, with a much smaller computational complexity, generates the highest PSNRs in most cases. Its gain on Barbara, rich in geometrical texture, is particularly significant. An example illustrated in Fig. 3 shows that the SSMS denoising restores some very fine patterns.

<sup>1</sup>The experimental results are available at <http://www.cmap.polytechnique.fr/~yu/research/ASIFT/demo.html>.

$\sigma/\text{PSNR}$	Lena		Barbara		House		Boat	
5/34.15	37.35	38.49	36.95	37.79	37.85	38.65	35.15	36.97
	38.60	<b>38.62</b>	38.08	<b>38.73</b>	39.37	<b>39.51</b>	<b>37.22</b>	37.09
10/28.13	34.86	35.61	34.28	34.03	35.39	35.35	32.65	33.58
	35.47	<b>35.63</b>	34.42	<b>35.11</b>	35.98	<b>36.13</b>	33.64	<b>33.70</b>
20/22.11	31.47	<b>32.66</b>	30.10	30.32	32.45	32.39	29.25	30.38
	32.38	32.30	30.83	<b>31.25</b>	<b>33.20</b>	32.77	30.36	<b>30.40</b>

**Table 1.** Denoising performance (in PSNR). Top left: NLmeans [4]. Top right: GSM [17]. Bottom left: learned overcomplete dictionary [6]. Bottom right: the proposed SSMS denoising. In each set the best result is in bold.



**Fig. 3.** Denoising. From left to right: a cropped region from Lena, noisy version ( $\sigma = 10$ ), the proposed SSMS denoising.

#### 4.1.2. Gaussian colored noise removal — deblurring

In these experiments, the clean image  $f$  is blurred and contaminated by a Gaussian white noise  $w$

$$y[n] = (f * h)[n] + w[n] \text{ for } 1 \leq n \leq N, \quad (11)$$

where  $y$  is the observed image and  $h$  is the blurring kernel that is assumed to be known. Deblurring aims at estimating  $f$  from  $y$ .

The deblurring can be casted to a denoising problem by first making a pseudo-inverse of  $h$  in the Fourier domain with a Wiener filter, and then removing the Gaussian colored noise amplified by the Wiener filtering [16].

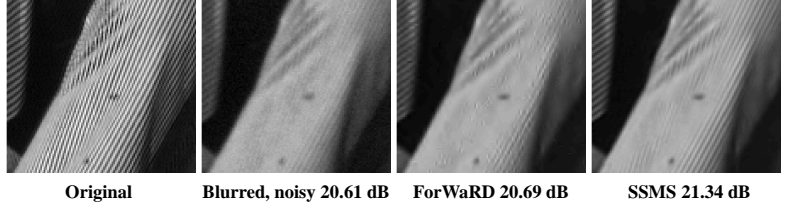
The SSMS deblurring is compared with the popular ForWaRD algorithm [16]. The same Wiener filter is applied in both methods. While the proposed SSMS deblurring proceeds the denoising with the SSMS model, ForWaRD denoises by translation-invariant wavelet shrinkage. A Monte Carlo procedure is applied for both methods to estimate the Gaussian colored noise variance on each coefficient. As summarized in Table 2, the SSMS deblurring improves on average about 0.7 dB in PSNR over ForWaRD. An example illustrated in Fig. 4 shows that the SSMS deblurring removes more effectively the colored noise and better restores fine patterns. Since the two methods apply the same Wiener filter, they achieve the same amount of deblurring, which will be further improved by the super-resolution deblurring described in Section 4.2.2.

$\sigma_b/\sigma_n$	Lena		Cameram.		House		Boat	
1 / 5	33.13	33.95	27.18	27.78	32.35	33.13	29.98	30.66
	33.18	<b>34.21</b>	27.09	<b>27.96</b>	32.31	<b>33.65</b>	30.22	<b>30.83</b>
2 / 5	31.17	31.95	24.51	24.63	30.31	30.87	27.84	28.22
	31.31	<b>32.11</b>	25.01	<b>25.02</b>	30.99	<b>31.95</b>	28.23	<b>28.38</b>

**Table 2.** Deblurring performance (in PSNR). The left column specifies the std of the Gaussian blur kernel  $\sigma_b$  and of the Gaussian white noise  $\sigma_n$ . Top left: ForWaRD [16]. Top right: the proposed SSMS deblurring (Sec 4.1.2). Bottom left: DSD [11]. Bottom right: the proposed SSMS super-resolution deblurring (Sec 4.2.2). In each set the best result is in bold.

## 4.2. Inverse Problems

Image enhancement often requires to solve an inverse problem, which amounts to estimate an image  $f$  in a space of larger dimension than



**Fig. 4.** Deblurring. From left to right: a cropped region from Barbara, blurred and noisy version, ForWaRD [16], the proposed SSMS deblurring. that of the degraded image  $y$  obtained through a linear operator  $U$ , and further contaminated by an additive noise  $w$ :

$$y[n] = Uf[n] + w[n] \text{ for } 1 \leq n \leq N. \quad (12)$$

The deblurring addressed in Section 4.1.2 is a classical inverse problem, where  $U$  is a convolution operator.

Given a dictionary  $\mathcal{D} = \{\phi_m\}_{m \in \Gamma}$  that provides a sparse representation for the original image  $f$  (1), the degraded image  $y$  is sparsely represented in the transformed dictionary  $\mathcal{D}_U = \{U\phi_m\}_{m \in \Gamma}$ :

$$y = y_\Lambda + w' \text{ with } y_\Lambda = \sum_{m \in \Lambda} a[m]U\phi_m, \quad (13)$$

where  $w' = U w_\Lambda + w$  and  $\|w'\|^2 \ll \|y\|^2$ . The sparse super-resolution approach [7, 11] seeks to estimate the coefficients  $\tilde{a}$  and the support  $\tilde{\Lambda} = \{m : \tilde{a}[m] \neq 0\}$  of  $f$  from the degraded signal  $y$ :

$$\tilde{a} = \arg \min_a \|y - \sum_{m \in \Gamma} a[m]U\phi_m\|^2 + \lambda \|a\|_1. \quad (14)$$

The resulting sparse super-resolution estimate of  $f$  is

$$\tilde{f} = \sum_{m \in \tilde{\Lambda}} \tilde{a}[m]\phi_m. \quad (15)$$

In addition to the sparse representation requirement on  $\mathcal{D}$ , the following conditions, not addressed in [7, 11], are necessary for a high-quality super-resolution estimate  $\tilde{f}$  [15]:

- **Recoverability.**  $\|U\phi_m\|^2 \gg 0, \forall m \in \Lambda$ .

- **Stability.**  $\mathcal{D}_U = \{U\phi_m\}_{m \in \Gamma}$  should be *incoherent* enough.

While the first condition suggests the *recoverability* of the support  $\Lambda$  from  $y$ , the second condition implies the *stability* of the support estimate. It has recently been shown that imposing structure in the estimation essentially stabilizes the estimate [2, 15].

The SSMS model enjoys the advantage of being highly structured. Instead of solving (14), SSMS calculates for each  $\mathcal{B}^k \in \mathcal{D}$

$$\tilde{a}^k = \arg \min_a e^k = \arg \min_a \|y - \sum_{m=1}^M a[m]U\phi_m^k\|^2 + \lambda \|a\|_1. \quad (16)$$

The best basis is obtained with  $k_0 = \arg \min_{k=1, \dots, K} e^k$ . (17)

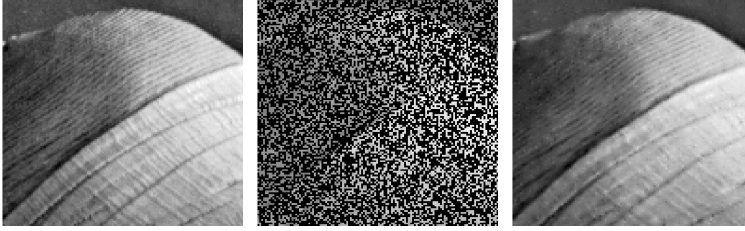
The sparse super-resolution (14) and (15) is then calculated with the best basis  $\mathcal{B}^{k_0}$ .

#### 4.2.1. Super-resolution inpainting

Inpainting aims at estimating  $f$  from  $y$  in (12), where  $U$  is a binary random mask on the image and  $w$  is typically neglected. The SSMS inpainting is calculated with patches of size  $8 \times 8$ ,  $M = 8$  and  $K = 18$ . The  $l^1$  minimization is calculated via a weighted Lasso with the LARS algorithm [12]. Table 3 summarizes the SSMS inpainting results, in comparison with the EM [9] and the MCA [7] algorithms that utilize a dictionary comprised of a curvelet frame and a local DCT, as well as with [10] that utilizes a local DCT. The results of MCA and EM are produced with the authors' softwares [7, 9]. [10] is a degenerated case of SSMS consisting of only one basis that is the DCT. The SSMS inpainting generates the highest PSNRs in all cases. When the masking percentage becomes important at 60%, the highly structured SSMS model significantly outperforms the second best algorithm by more than 2 dB in PSNR. As illustrated in Fig 5, the SSMS inpainting accurately restores the image geometry from a very heavily masked image.

masked/PSNR	Lena		Barbara		House		Boat	
30%/10.14	40.32	33.40	38.56	32.46	41.95	33.57	37.33	31.34
	38.53	<b>40.76</b>	38.95	<b>40.70</b>	40.67	<b>43.00</b>	36.17	<b>37.69</b>
60%/7.11	33.43	29.91	29.77	28.52	34.10	29.56	30.40	27.58
	33.95	<b>36.02</b>	32.33	<b>34.50</b>	35.23	<b>37.41</b>	31.00	<b>32.53</b>

**Table 3.** Inpainting performance (in PSNR). Top left: [10]. Top right: EM [9]. Bottom left: MCA [7]. Bottom right: the proposed SSMS inpainting. In each set the best result is in bold.



**Fig. 5.** Inpainting. From left to right: a cropped region from Lena, 60% randomly masked version, the proposed SSMS inpainting.

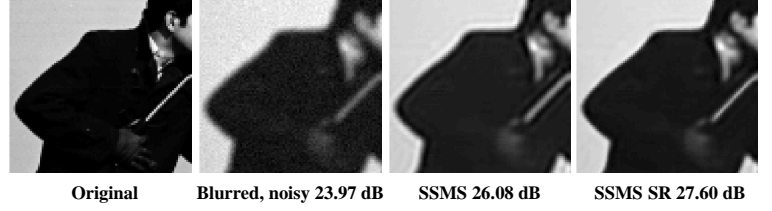
#### 4.2.2. Super-resolution deblurring

Super-resolution deblurring has the capacity to restore image patterns at frequencies eliminated by the convolution operator  $U$ , which is impossible with the Wiener filtering approach described in Section 4.1.2. The SSMS super-resolution deblurring is calculated with a hierarchy of two families of orthogonal bases.

The directional bases introduced in Section 3.3.1 do not have super-resolution deblurring ability since they do not satisfy the *recoverability* condition described above. Indeed these oscillatory atoms, as illustrated in Fig. 2, have Dirac supports in Fourier and can be therefore killed after the convolution. Following the hierarchical sparse decomposition scheme decoupling orientation and position introduced in [19], a family of orthogonal bases that specifies the edge position in the patch is cascaded with each directional basis of the corresponding orientation. The atoms in the position bases are localized around the underlying position, and thus have spread spectrum in Fourier as required by the *recoverability* condition. The position bases are computed with the PCA procedure in the same way as described in Section 3.3.1, for each basis the training patches being extracted at a fixed position from the synthetical contour images. Two basis selections (17) are cascaded in the hierarchy: the first selects the best direction basis and, given the best orientation, the second selects the best position basis. The sparse super-resolution deblurring is then calculated with the best position basis in the selected direction.

The SSMS super-resolution deblurring is conducted with patches of size  $12 \times 12$ ,  $M = 20$  and  $K = 18$ . The  $l^1$  minimization is calculated via a weighted Lasso with the LARS algorithm [12]. The patch boundary issue is treated in the same way as in [11]. Table 2 summarizes the performance of the SSMS super-resolution deblurring, in comparison with ForWaRD, the SSMS deblurring (without super-resolution), and the direct sparse deconvolution (DSD) [11].<sup>1</sup> The SSMS super-resolution deblurring generates the highest PSNRs in all cases, followed by the SSMS deblurring. As its gain with respect to the latter lies essentially on edges that take a small proportion in the image, the PSNR improvement over the whole image is not significant. However, the visual benefit is substantial. As illustrated in

Fig. 6, the super-resolution deblurring restores sharper edges than the SSMS deblurring without super-resolution described in 4.1.2.



**Fig. 6.** Deblurring. From left to right: a cropped region from Cameraman, blurred and noisy version, the proposed SSMS deblurring (without super-resolution) and the proposed SSMS super-resolution deblurring.

To conclude, let us mention that we are currently extending this work to color image modeling and enhancement.

**Acknowledgements:** The authors would like to thank Michael Elad, and Jean-Michel Morel for the inspiring discussions.

## 5. REFERENCES

- [1] M. Aharon, M. Elad, and A. Bruckstein. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Trans on Signal Processing*, 54(11):4311–4322, 2006.
- [2] R.G. Baraniuk, V. Cevher, M.F. Duarte, and C. Hegde. Model-based compressive sensing. *Submitted to IEEE Trans on Information Theory*, 2008.
- [3] A. Barron, L. Birgé, and P. Massart. Risk bounds for model selection via penalization. *Prob Theo and Relat Fields*, 113(3), 1999.
- [4] A. Buades, B. Coll, and J.M. Morel. A review of image denoising algorithms, with a new one. *SIAM MMS*, 4(2):490–530, 2006.
- [5] S.S. Chen, D.L. Donoho, and M.A. Saunders. Atomic decomposition by basis pursuit. *SIAM J on Sci Comput*, 20:33–61, 1999.
- [6] M. Elad and M. Aharon. Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Trans on Image Processing*, 15(12):3736–3745, 2006.
- [7] M. Elad, J.L. Starck, P. Querre, and DL Donoho. Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA). *ACHA*, 19(3):340–358, 2005.
- [8] Y.C. Eldar, P. Kuppinger, and H. Bölcskei. Compressed sensing of block-sparse signals: Uncertainty relations and efficient recovery. *Submitted to IEEE Trans on Signal Processing*, 2009.
- [9] M.J. Fadili, J.L. Starck, and F. Murtagh. Inpainting and zooming using sparse representations. *The Computer J*, 52(1):64, 2009.
- [10] O.G. Guleryuz. Nonlinear approximation based image recovery using adaptive sparse reconstructions and iterated denoising-Part II: Adaptive algorithms. *IEEE Trans on Image Processing*, 15, 2006.
- [11] Y. Lou, A. Bertozzi, and S. Soatto. Direct sparse deblurring. *Technical Report, CAM-UCLA*, 2009.
- [12] J. Mairal, F. Bach, J. Ponce and G. Sapiro. Online learning for matrix factorization and sparse coding. *To appear in Journal of Machine Learning Research*, 2009.
- [13] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman. Non-local sparse models for image restoration. In *Proc. ICCV*, 2009.
- [14] J. Mairal, M. Elad, and G. Sapiro. Sparse representation for color image restoration. *IEEE Trans on Image Processing*, 17(1), 2008.
- [15] S. Mallat and G. Yu. Super-resolution with mixing sparse estimators. *Submitted to IEEE Trans Image Processing*, 2009.
- [16] R. Neelamani, H. Choi, and R. Baraniuk. Forward: Fourier-wavelet regularized deconvolution for ill-conditioned systems. *IEEE Trans on signal processing*, 52(2):418–433, 2004.
- [17] J. Portilla, V. Strela, M.J. Wainwright, and E.P. Simoncelli. Image denoising using scale mixtures of Gaussians in the wavelet domain. *IEEE Trans on Image Processing*, 12(11):1338–1351, 2003.
- [18] J.A. Tropp and A.C. Gilbert. Signal recovery from random measurements via orthogonal matching pursuit. *IEEE Trans on Information Theory*, 53(12):4655–4666, 2007.
- [19] G. Yu and S. Mallat. Sparse super-resolution with space matching pursuits. In *Proc. SPARS’09, Saint-Malo*, 2009.

<sup>1</sup>We thank very much Yifei Lou for kindly making the DSD experiments.